AMENDMENTS TO THE SPECIFICATION

Please replace the paragraph beginning at page 12, line 19, which starts with "Figure 1A," with the following amended paragraph:

Figure 1A illustrates a GFP gene targeting system. The artificial gene target (A658) consisted of a GFP gene mutated by a 35 basepair insertion which includes a stop codon and a recognition site for the I-SceI endonuclease (Sce) (5' TAGGGATAACAGGGTAAT 3', SEQ ID NO: 1) at basepair 327 of the coding sequence. The GFP gene was driven by a hybrid cytomegalovirus enhancer/chicken β3-actin promoter ("CMV/CBA" or "CBA"). The GFP gene was part of bicistronic transcript in which an internal ribosomal entry site ("IRES") allowed translation of the human CD8α gene ("CD8"). The bicistronic message contained a Woodchuck post-trancriptional regulatory element ("WPRE") to increase messenger RNA levels (Zufferey et al., 1999, J Virol, 73:2886-92). Finally, the locus contained a gene with the phosphoglycerate kinase promoter ("PGK") driving the neomycin phosphotransferase gene (NEO) to allow selection by the antibiotic G418. The repair substrates RS2100 and RS2700 are also depicted. They consisted of a GFP gene that has been truncated at basepair 37 of the coding sequence and thus were missing the initiation codon ("truncGFP"). The truncated GFP gene was followed by the IRES-CD8 for RS2100 or IRES-CD8-WPRE for RS2700 as in A658. The A658 gene target was introduced into 293 cells by electroporating 2 x 10^6 cells with 10 μg of supercoiled A658 plasmid DNA. Cells were selected in 500 μg/ml G418 for two weeks. Monoclonal cell lines were made by picking individual colonies and identifying those with high surface CD8 expression by staining with phycoerythrin-conjugated anti-CD8 antibody (BD Biosciences, San Jose, CA) (293 cells normally do not express CD8). Polyclonal cell lines were made by purifying a population of cells using Miltenyi anti-CD8 microbeads and a MACS minicolumn (Miltenyi Biotec, Auburn, CA). Gene targeting was measured by transfecting 293/A658 cells with RS2100 with or without a Sce expression plasmid along with a control plasmid (pON405) to determine the transfection efficiency. Applicants used three different promoters to drive Sce expression: PGK, cytomegalovirus ("CMV"), and CBA. The cells were then incubated for 3 days and the percentage of GFP positive cells measured by flow cytometry using a FACScan

(BD Biosciences, San Jose, CA). The gene targeting rate was determined by normalizing the measured percentage of GFP positive cells to the transfection efficiency.

Please replace the paragraph beginning at page 15, line 18, which starts with "Figure 3A", with the following amended paragraph:

Figure 3A shows schematics of the chimeric nucleases and chimeric nuclease targets. The gene targets were identical to A658 except that additional sequence, which expands the insertion, have been inserted into the GFP gene adjacent to the Sce recognition site ("Sce site"). In QQR8 and QQR6 inverted repeats of a the binding site ("QQR site") for the QQR zinc finger triplet (5' GGGGAAGAA 3', SEQ ID NO: 2) were inserted with either a 6 bp, "6," (QQR6) or 8 bp spacer, "8," (QQR8). In QQRZIF6, a binding site for the Zif268 triplet finger ("Zif Site") (5' GCGTGGTCG 3', SEQ ID NO: 3) was inserted in an inverted orientation to a QQR site with a 6 bp spacer ("6") between the sites. Polyclonal 293 cell lines were made from QQR8, QQR6, and QQRZIF6 as described in Figure 1. The chimeric nucleases were driven by the CMV promoter, "CMV." Each have a standard initiation codon "ATG" followed by a nuclear localization signal, "N," at the amino-terminus. The triplet zinc finger domain, either "QQR" for the QQR zinc finger triplet (Shi et al., 1995, Science, 268:282-284) or "Zif" for the Zif268 triplet (Wolfe et al., 2001, Structure (Camb), 9:717-23) follow the nuclear localization signal. There is then a variable amino acid linker, 18 amino acids ("L18") in CMV-QQR-L18-Fn, zero amino acids ("L0") in CMV-QQR-L0-Fn, or three amino acids ("L3") in CMV-ZIF-L3-Fn before the endonuclease domain of the FokI restriction enzyme ("Nuclease" or "Fn") (Chandrasegaran et al., 1999, Biol Chem, 380:841-8). CMV-QQR-L18-Fn and CMV-QQR-LO-Fn were cloned from previously characterized fusion proteins (Smith et al., 2000, Nucleic Acids Res, 28:3361-9) while CMV-ZIF-L3-Fn is novel.

Please replace the paragraph beginning at page 16, line 14, with the following amended paragraph:

Figure 4 demonstrates the sequence (SEQ ID NO: 16) of the human β-globin gene surrounding the codon mutated (in red) to cause sickle cell anemia. Depicted are two pairs of potential chimeric nucleases (HBGZF1 and HBGZF2; HBGZF3 and HBGZF4). The binding sites for the chimeric nucleases are highlighted by being in capital letters.

Please replace the paragraph beginning at page 16, line 18, with the following amended paragraph:

Figure 5 demonstrates the binding site for HBGZF1 (SEQ ID NO: 17) and the zinc finger domains (SEQ ID NOs: 18-23) that recognize each triplet using the single letter code.

Please replace the paragraph beginning at page 16, line 20, with the following amended paragraph:

Figure 6 shows the results of gene targeting with HBGZF1 and the GFP gene target (SEQ ID NO: 24) containing the artificial hybrid HBGZF1/Zif268 binding site.

Please replace the paragraph beginning at page 16, line 22, with the following amended paragraph:

Figure 7 shows the design and target site for HBGZF4. SEQ ID NOs: 25-34.

Please replace the paragraph beginning at page 16, line 23, with the following amended paragraph:

Figure 8 shows the results of gene targeting with HBGZF4 and the GFP gene target (SEQ ID NO: 35) containing the artificial hybrid HBGZF4/Zif268 binding site.

Please replace the paragraph beginning at page 16, line 25, with the following amended paragraph:

Figure 9A shows the structure of the human common γ-chain and the location of mutations in the gene that lead to SCID, derived from Notarangelo et al, 2002. Figure 9B shows the sequence of exon 5 and the proposed binding sites for chimeric nucleases HCGCZFl and HCGCZF2. SEQ ID NOs: 36-37.

Please replace the paragraph beginning at page 17, line 3, with the following amended paragraph:

Figure 10 shows the binding site for HCGCZF2 and the structure of HCGCZF2 using the amino acids for zinc fingers 1-3 deduced from the zinc-finger code from Sera and Uranga (2002). SEQ ID NOs.: 38-48.

Please replace the paragraph beginning at page 17, line 6, with the following amended paragraph:

Figure 11 shows the results of gene targeting with HBGZF2 and the GFP gene target (SEQ ID NO: 49) containing the artificial hybrid HCGCZF2/Zif268 binding site.

Please replace the paragraph beginning at page 17, line 8, with the following amended paragraph:

Figure 12: Gene Targeting with GFP chimeric nucleases. A) The sequence (SEQ ID NO: 50) of the target sequence in GFP gene and a schematic representation of chimeric nucleases designed to cleave the GFP gene. The GFP chimeric nuclease target site lies just 5' to the insertion of the I-SceI recognition site ("Sce site"). B) Rate of gene targeting in 293 cells after co-transfection of the indicated nuclease with the repair substrate A767 described in example 1.

Please replace the paragraph beginning at page 17, line 14, with the following amended paragraph:

Figure 13: Gene Targeting using CD8 Chimeric Nucleases. A) Target sequence (SEQ ID NOs: 51 and 52) within human CD8α gene for chimeric nucleases. B) Flow cytometry plots after transfecting 293/1104 cells with the CD8 Knockout Plasmid alone (5% CD8 negative cells) or with the CD8 Knockout Plasmid plus the CD8 chimeric nucleases (20% CD8 negative). The measurement of CD8 expression was done after selecting for puromycin resistant colonies and by staining with phycoerytherin conjugated α-CD8 monoclonal antibody.

Please replace the paragraph beginning at page 23, line 19, which starts with "In certain embodiment", with the following amended paragraph:

In certain embodiments, the disclosure provides chimeric nucleases that are particularly effective for use in gene targeting methods. In certain gene targeting protocols, it may be desirable to cause a DNA cleavage near or at the target sequence while also keeping a limit on the number of cleavages that occur in other portions of the genome. Accordingly, it may be desirable to employ a chimeric nuclease or cooperatively acting set of chimeric nucleases that have a recognition sequence occurring rarely or uniquely in the genome to be altered. As a general principle, the larger the recognition sequence, the fewer times such sequence is likely to occur in the genome to be altered. A simple probability calculation suggests that a recognition sequence having *n* defined nucleotides will occur with a probability of one in 4^n nucleotides. According to this simplified predictive model, a recognition sequence of 11 nucleotides is most likely to occur once in the genome of an *Escherichia coli* bacterium (genome size of roughly 4.5 million bases). The human genome is estimated at a size of 3 billion base pairs, and so a chimeric nuclease having a 16 nucleotide recognition sequence is most likely to recognize only a single sequence. The simple statistical model may be adjusted to account for factors such as GC bias, repeat sequences, and heterogeneity in the target organism (e.g. humans vary by roughly 1% from each other, and such additional variation could be taken into account). Furthermore,

recognition sequences may be assessed by searching for actual occurrences of the sequence in published genomic sequence of the target organism. A recognition sequence may be contiguous (an uninterrupted string of defined nucleotides, *e.g.*, 5'-GATGTTGCT-3', SEQ ID NO: 4) or non-contiguous (interrupted by non-defined nucleotides, *e.g.*, 5'-GATG... N_6...TTGCT-3', SEQ ID NO: 5), and in either case the frequency of occurrence can be estimated in the same way.

Please replace the paragraph beginning at page 27, line 14, which starts with "The spectrum of possible recognition sequences", with the following amended paragraph:

The spectrum of possible recognition sequences may be compared against the region that is in workable proximity of the target sequence. To be effective for gene targeting, a DNA binding domain need that will be coupled to a cleavage domain need only bind so as to permit cleavage within a workable proximity of the target sequence. A workable proximity is any distance that still facilitates the gene targeting. In certain embodiments, a workable proximity is within at least 500 base pairs of the most distal target sequence to be changed, preferably within 200 base pairs and most preferably within 100 or 50 base pairs of the most distal target sequence to be changed. Optionally, the DNA binding domain overlaps the target sequence. Given that a target sequence is defined herein as the sequence to be altered, a target sequence may stretch over a plurality of nucleotides. In such situation, a DNA binding domain may, of course, bind within the target sequence, and the term "workable proximity" is intended to encompass this scenario. Selecting a DNA binding site may also involve evaluating the likelihood that a particular recognition sequence occurs elsewhere in the genome, and methods for doing so are described above. As described in the examples below, a variety of chimeric zinc finger nucleases may be generated. Families of such proteins will tend to bind certain consensus sequences, such as the 5' GNNGNNGNN 3' (SEQ ID NO: 6) sequence, dimers of which could recognize 5' NNCNNCNNC NNNNNN GNNGNNGNN 3' (SEQ ID NO: 7) (predicted to occur roughly once per 4096 bases in a genome). Accordingly, a region to be targeted may be scanned for a workable consensus recognition sequence, and then a zinc finger that recognizes the specific sequence may be designed. By searching a target region for a consensus sequence and then designing a suitable specified chimeric nuclease. The workability of this approach is demonstrated in example 3, and such techniques may be applied to other zinc finger nucleases and other chimeric nucleases generally.

Please replace the paragraph beginning at page 47, line 14, which starts with "In the GFP gene targeting system", with the following amended paragraph:

In the GFP gene targeting system the introduction of a DSB stimulated GT by >2000-fold and the absolute rate of gene targeting reached 3-5% when conditions were optimized. Such a system, however, depended on the prior introduction of a Sce site into the target gene and therefore can not be used for endogenous genes. To stimulate gene targeting at endogenous genes, a method to create sequence specific DSBs in those genes needs to be developed. Chimeric nucleases have such potential (Chandrasegaran et al., 1999, Biol Chem, 380:841-8). Chimeric nucleases--fusions between zinc finger binding DNA binding domains and the endonuclease domain of the FokI restriction enzyme ("Fn")--can site-specifically cleave naked DNA *in vitro* (Chandrasegaran et al., 1999, Biol Chem, 380:841-8), extra-chromosomal DNA in Xenopus oocytes (Bibikova et al., 2001, Mol Cell Biol, 21:289-97) and chromosomal DNA in Drosophila (Bibikova, et al., 2002, Genetics, 161:1169-75). Applicants decided to try to extend this methodology to stimulate gene targeting in human somatic cells (Figure 3). Figure 3A shows the structure of the expression plasmids and target sites for the chimeric nuclease experiments. Applicants designed three different chimeric nucleases, each driven by the CMV promoter and containing a nuclear localization signal at their amino-termini (Figure 3A). In two constructs (CMV-QQR-L18-Fn and CMV-QQR-L0-Fn) the DNA binding specificity was conferred by the artificial QQR three zinc finger domain that binds with nanomolar affinity to the sequence 5' GGGGAAGAA 3' (SEQ ID NO: 8) (Shi et al., 1995, Science, 268:282-284). These two constructs differed in the length of the amino acid linker between the zinc fingers and the Fn domain. The amino acid linker was 18 amino acids in CMV-QQR-L18-Fn while in CMV-QQR-L0-Fn there was no amino acid linker. CMV-ZIF-L3-Fn fused the three zinc fingers from Zif268 to the Fn domain with a 3 amino acid linker between the two domains. The Zif268 zinc finger domain recognizes the sequence 5' GCGTGGGCG 3' (SEQ ID NO: 9) with sub-nanomolar affinity (Elrod-Erickson et al., 1999, J Biol Chem, 274:19281-5). Applicants constructed three cell lines (293/QQR8, 293/QQR6, 293/QQRZIF6) with corresponding gene targets (QQR8, QQR6, and QQRZIF6). QQR8 and QQR6 have inverted repeats of the QQR binding site inserted next to the Sce recognition site (Figure 3A). QQR8 and QQR6 differed in that the repeats are separated by 8 bp in QQR8 and 6 bp in QQR6. Prior work has shown that purified QQR-Fn protein without an amino acid linker (equivalent to CMV-QQR-L0-Fn) cuts

DNA most efficiently when the inverted DNA binding sites are separated by 6 bp while purified QQR-Fn protein with an 18 amino acid linker (equivalent to CMV-QQR-L18-Fn) cuts DNA when the binding sites are separated by either 6 or 8 bp (Bibikova, et al., 2002, Genetics, 161:1169-75; Smith et al., 2000, Nucleic Acids Res, 28:3361-9). In all of the chimeric nuclease gene targeting experiments, Applicants co-transfected the chimeric nuclease with the repair substrate RS2700 (Figure 1A).

Please replace the paragraph beginning at page 54, line 18, which starts with "In Example 2", with the following amended paragraph:

In example 2, applicants demonstrated the design of chimeric nucleases (zinc finger nucleases in this example) to cleave at sequences derived from endogenous genes (the β-globin gene and the common γ-chain gene). In particular applicants were 100% successful at designing chimeric nucleases to cleave at target sites with the following consensus sequence: 5' GNNGNNGNN 3' (SEQ ID NO: 6) where G represents guanine and N represents any nucleotide. Applicants expected, therefore, that one could empirically design a pair of zinc finger nucleases to stimulate gene targeting in a natural gene if that gene contained an inverted repeat of the above consensus sequence with the repeats separated by 6 nucleotides (e.g. 5' NNCNNCNNC NNNNNN GNNGNNGNN 3', SEQ ID NO: 7). Both the GFP gene and the human CD8α gene contain such a sequence.

Please replace the paragraph beginning at page 55, line 2, which starts with "By searching the sequence of GFP," with the following amended paragraph:

By searching the sequence of GFP, following sequence was identified: 5' ACC ATC TTC TTCAAG GAC GAC GGC 3' (SEQ ID NO: 10). This sequence fits the inverted repeat consensus sequence described above. The sequence goes from bp 292-315 of the coding sequence of the gene. Using, in part, guidance from Liu et al. (2002) applicants made GFP-CN1 to recognize the target sequence 5' GAA GAT GGT 3' (SEQ ID NO: 11) and GFP-CN2 to recognize the target sequence 5' GAC GAC GGC 3' (SEQ ID NO: 12). Applicants used the chimeric nuclease backbone from Zif-Fn to make each of these and tested the GFP-CNs using the GFP gene targeting system described in example 1 above. The target binding site for the GFP-CNs lies adjacent to the insertion of the I-SceI recognition site into the GFP gene. Briefly, in these cells the mutated GFP gene is integrated as a single copy into the genome of 293 cells.

The GFP-CNs were able to stimulate gene targeting when co-transfected with a repair substrate by 1000-fold (Figure 12). The GFP-CNs were not as efficient as I-SceI at stimulating gene targeting (Figure 12).

Please replace the paragraph beginning at page 55, line 19, which starts with "In the GFP gene targeting system," with the following amended paragraph:

In the GFP gene targeting system we express the human CD8α gene off a bicistronic transcript that includes the mutated GFP gene. In cell line 293/1004, for example, 95% of the cells are CD8 positive. Applicants found the following sequence 5' GGCGCCCAC CATCGC GTCGCAGCC 3' (SEQ ID NO: 13) that spans base pair 441-468 of the human CD8α gene and fits the inverted repeat consensus described above. Applicants constructed CD8CN1 to recognize 5' GTGGGCGCC 3' (SEQ ID NO: 14) and CD8CN2 to recognize 5' GTCGCAGCC 3' (SEQ ID NO: 15). Applicants also constructed a CD8 cDNA knock-out plasmid in which a puromycin resistance cassette is flanked by 440 bases of 5' homology and 220 bases of 3' homology to the CD8 gene (called "CD8 Knockout Plasmid"). Applicants then transfected the CD8 Knockout Plasmid with and without the CD8 chimeric nucleases into cell line 293/1104 and measured the percentage of CD8 positive cells in a population of cells after puromycin selection (Figure 13). Transfecting the CD8 Knockout Plasmid did not change the number of CD8 negative cells from the parent population (5% CD8 negative in both) as expected. After co-transfection of the CD8 chimeric nucleases with the CD8 Knockout Plasmid, over 20% of the cells were now CD8 negative. This shows that chimeric nucleases can stimulate gene targeting in the CD8α cDNA by stimulating the insertion of the puromycin knockout plasmid into the gene.